

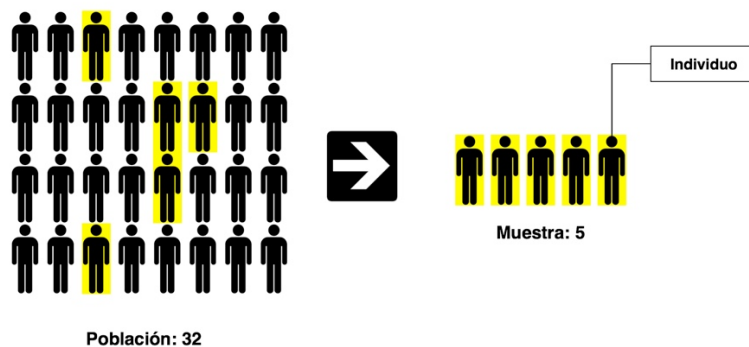
# Estadística

## Población y muestra. Variables estadísticas

**Población** es el conjunto de elementos sobre que se realiza un estudio estadístico.

**Muestra** es la parte de la población que estudiamos. Su tamaño es el número de elementos que lo forman.

**Individuo** es cada uno de los elementos de la población o muestra.



### EJEMPLO

Se va a realizar un estudio estadístico sobre el porcentaje de personas casadas en una localidad de 122594 habitantes. Para ello se eligen 2325 habitantes y se extienden las conclusiones a toda la población:

Variable estadística	Si una persona está casada o no.
Población	Los 122594 habitantes de la localidad.
Muestra	Los 2325 habitantes elegidos aleatoriamente.
Tamaño de la muestra	2325 personas.
Individuo	Cada persona a la que se pregunta.

Una **variable estadística** es cada una de las propiedades o características que podemos estudiar de una población o muestra.

Variables estadísticas	Cualitativas (raza, idioma)	
	Cuantitativas (valores numéricos)	Discretas Número finito de valores (número de cartas)
		Continuas Número infinito de valores (peso)

# Tablas de frecuencias

**Frecuencia absoluta  $f_i$**  de un dato es el número de veces que aparece en la muestra.

**Frecuencia relativa  $h_i$**  de un dato es el cociente de la frecuencia absoluta y el número total de datos,  $N$

**Frecuencia absoluta acumulada  $F_i$**  de un dato es la suma de todas las frecuencias absolutas de los valores menores o iguales que él.

**Frecuencia relativa acumulada  $H_i$**  de un dato es la suma de todas las frecuencias relativas de los valores menores o iguales que él.

Tras efectuar el recuento de datos, los valores de la variable y las frecuencias se organizan en una **tabla de frecuencias**.

## EJEMPLO

Organiza en una tabla las estaturas, en cm, de un grupo de niños.

130    128    141    139    137    126    135    136    134    131  
143    140    129    128    137    136    142    138    144    136

La **estatura** es una variable estadística **cuantitativa continua**: por tanto, podemos agrupar los datos en intervalos que llamaremos **clases**. La amplitud de cada intervalo viene dada por la fórmula:

$$\frac{Max - Min}{\sqrt{N}} \rightarrow \begin{cases} Max = \text{Dato con valor máximo} \\ Min = \text{Dato con valor mínimo} \\ N = \text{Número total de datos} \end{cases}$$

Para el problema que tenemos:

$$Max = 144$$

$$Min = 126$$

$$N = 20$$

$$\frac{144 - 126}{\sqrt{20}} = 4,02$$

Aproximando, podemos decir que la amplitud de cada intervalo será de 4cm.

Estatura	$f_i$	$h_i$	$F_i$	$H_i$
Intervalo 1 [126,130)	$f_1 = 4$	$h_1 = \frac{f_1}{N} = \frac{4}{20} = 0,2$	$F_1 = 4$	$H_1 = 0,2$
Intervalo 2 [130,134)	$f_2 = 2$	$h_2 = \frac{f_2}{N} = \frac{2}{20} = 0,1$	$F_2 = 4 + 2 = 6$	$H_2 = 0,2 + 0,1 = 0,3$
Intervalo 3 [134,138)	$f_3 = 7$	$h_3 = \frac{f_3}{N} = \frac{7}{20} = 0,35$	$F_3 = 6 + 7 = 13$	$H_3 = 0,3 + 0,35 = 0,65$
Intervalo 4 [138,142)	$f_4 = 4$	$h_4 = \frac{f_4}{N} = \frac{4}{20} = 0,2$	$F_4 = 13 + 4 = 17$	$H_4 = 0,65 + 0,2 = 0,85$
Intervalo 5 [142,146)	$f_5 = 3$	$h_5 = \frac{f_5}{N} = \frac{3}{20} = 0,15$	$F_5 = 17 + 3 = 20$	$H_5 = 0,85 + 0,15 = 1$

De nuevo la tabla, quitando cálculos:

Estatura	$f_i$	$h_i$	$F_i$	$H_i$
[126, 130)	4	0,2	4	0,2
[130, 134)	2	0,1	6	0,3
[134, 138)	7	0,35	13	0,65
[138, 142)	4	0,2	17	0,85
[142, 146)	3	0,15	20	1

La última fila de la frecuencia absoluta acumulada  $F_i$  debe ser igual a **N**

La última fila de la frecuencia relativa acumulada  $H_i$  debe ser igual a **1**